# Exhibit F

**AskF5**   Knowledge Centers   Resources

**My Support**

**Manual Chapter** : Managing Traffic with Rate Shaping

**Applies To:**

Show Versions ⊞

# Managing Traffic with Rate Shaping

### Introduction to rate shaping

The BIG-IP® system includes a feature called rate shaping. **Rate shaping** allows you to enforce a throughput policy on incoming traffic. Throughput policies are useful for prioritizing and restricting bandwidth on selected traffic patterns.

Rate shaping can be useful for an e-commerce site that has preferred clients. For example, the site might want to offer higher throughput for preferred customers, and lower throughput for other site traffic.

The rate shaping feature works by first queuing selected packets under a rate class, and then dequeuing the packets at the indicated rate and in the indicated order specified by the rate class. A **rate class** is a rate-shaping policy that defines throughput limitations and a packet scheduling method to be applied to all traffic handled by the rate class.

You configure rate shaping by creating one or more rate classes and then assigning the rate class to a packet filter or to a virtual server. You can also use the iRules® feature to instruct the BIG-IP system to apply a rate class to a particular connection.

You can apply a rate class specifically to traffic from a server to a client or from a client to a server. If you configure the rate class for traffic that is going to a client, the BIG-IP system does not apply the throughput policy to traffic destined for the server. Conversely, if you configure the rate class for traffic that is going to a server, the BIG-IP system does not apply the throughput policy to traffic destined for the client.

### About rate classes

A rate class defines the throughput limitations and packet scheduling method that you want the BIG-IP® system to apply to all traffic that the rate class handles. You assign rate classes to virtual servers and packet filter rules, as well as through iRules®.

If the same traffic is subject to rate classes that you have assigned from more than one location, the BIG-IP system applies the last-assigned rate class only. The BIG-IP system applies rate classes in the following order:

- The first rate class that the BIG-IP system assigns is from the last packet filter rule that matched the traffic and specified a rate class.
- The next rate class that the BIG-IP system assigns is from the virtual server; if the virtual server specifies a rate class, the rate class overrides any rate class that the packet filter selects.
- The last rate class assigned is from the iRule; if the iRule specifies a rate class, this rate class overrides any previously-selected rate class.

**Note:** *Rate classes cannot reside in partitions. Therefore, a user's ability to create and manage rate classes is defined by user role, rather than partition-access assignment.*

You can create a rate class using the BIG-IP Configuration utility. After you have created a rate class, you must assign it to a virtual server or a packet filter rule, or you must specify the rate class from within an iRule.

### Rate class name

The first setting you configure for a rate class is the rate class name. Rate class names are case-sensitive and might contain letters, numbers, and underscores (_) only. Reserved keywords are not allowed.

Each rate class that you define must have a unique name. This setting is required.

To specify a rate class name, locate the `Name` field on the New Rate Class screen and type a unique name for the rate class.

### Base rate

The **Base Rate** setting specifies the base throughput rate allowed for traffic that the rate class handles. The sum of the base rates of all child rate classes attached to a parent rate class, plus the base rate of the parent rate class, cannot exceed the ceiling of the parent rate class. For this reason, F5 Networks® recommends that you always set the base rate of a parent rate class to 0 (the default value).

You can specify the base rate in bits per second (bps), kilobits per second (Kbps), megabits per second (Mbps), or gigabits per second (Gbps). The default unit is bits per second. This setting is required.

**Note:** *These numbers are powers of 10, not powers of 2.*

### Ceiling rate

The **Ceiling Rate** setting specifies the absolute limit at which traffic is allowed to flow when bursting or borrowing. You can specify the ceiling in bits per second (bps), kilobits per second (Kbps), megabits per second (Mbps), or gigabits per second (Gbps). The default unit is bits per second.

If the rate class is a parent rate class, the value of the ceiling defines the maximum rate allowed for the sum of the base rates of all child rate classes attached to the parent rate class, plus the base rate of the parent rate class.

**Note:** *A child rate class can borrow from the ceiling of its parent rate class.*

### Burst size

You use the **Burst Size** setting when you want to allow the rate of traffic flow that a rate class controls to exceed the base rate. Exceeding the base rate is known as *bursting*. When you configure a rate class to allow bursting (by specifying a value other than 0), the BIG-IP® system saves any unused bandwidth and uses that bandwidth later to enable the rate of traffic flow to temporarily exceed the base rate. Specifying a burst size is useful for smoothing out traffic patterns that tend to fluctuate or exceed the base rate, such as HTTP traffic.

The value of the **Burst Size** setting defines the maximum number of bytes that you want to allow for bursting. Thus, if you set the burst size to 5,000 bytes, and the rate of traffic flow exceeds the base rate by 1,000 bytes per second, then the BIG-IP system allows the traffic to burst for a maximum of five seconds.

When you specify a burst size, the BIG-IP system creates a burst reservoir of that size. A burst reservoir stores bandwidth that the BIG-IP system uses for bursting later. The burst reservoir becomes depleted as the rate of traffic flow exceeds the base rate, and is replenished as the rate of traffic falls below the base rate. The Burst Size value that you configure in a rate class thus represents:

- The maximum number of bytes that the rate class transmits when the traffic-flow rate exceeds the base rate
- The maximum number of bytes that the BIG-IP system can replenish into the burst reservoir
- The amount of bandwidth initially available for bursting beyond the base rate

The burst size is measured in bytes. For example, a value of either 10000 or 10K equals 10,000 bytes. The default value is 0.

### Depleting the burst reservoir

When the rate of traffic flow exceeds the base rate, the BIG-IP® system automatically depletes the burst reservoir, at a rate determined by the number of bytes per second that the traffic flow exceeds the base rate.

Continuing with our previous example in which traffic flow exceeds the base rate by 1,000 bytes per second, if the traffic-flow rate only exceeds the base rate for two seconds, then 2,000 bytes are depleted from the burst size and the maximum bytes available for bursting decreases to 3,000.

**Note:** *In some cases, a rate class can borrow bandwidth from the burst reservoir of its parent class.*

### Replenishing a burst reservoir

When the rate of traffic flow falls below the base rate, the BIG-IP® system stores the unused bandwidth (that is, the difference between the base rate and the actual traffic-flow rate) in the burst reservoir. Later, the BIG-IP system uses this bandwidth when traffic flow exceeds the base rate. Thus, the BIG-IP system replenishes the burst reservoir whenever it becomes depleted due to traffic flow exceeding the base rate.

The size of the burst reservoir cannot exceed the specified burst size. For this reason, the BIG-IP system replenishes the reservoir with unused bandwidth only until the reservoir reaches the amount specified by the **Burst Size** setting. Thus, if the burst size is set to 5,000, then the BIG-IP system can store only 5,000 bytes of unused bandwidth for later use when the rate of traffic flow exceeds the base rate.

**Note:** *Specifying a burst size does not allow the rate class to exceed its ceiling.*

### About specifying a non-zero burst size

This example illustrates the behavior of the BIG-IP® system when you set the **Burst Size** setting to a value other than 0.

This example shows throughput rates in units of bytes-per-second instead of the default bits-per-second. This is only to simplify the example. You can derive bytes-per-second from bits-per-second by dividing the bits-per-second amount by 8.

Suppose you configure the rate class settings with these values:

- Base rate: 1,000 bytes per second
- Ceiling rate: 4,000 bytes per second
- Burst size: 5,000 bytes

Consider the following scenario:

**If traffic is currently flowing at 800 bytes per second**
No bursting is necessary because the rate of traffic flow is below the base rate defined in the rate class. Because the traffic is flowing at 200 bytes per second less than the base rate, the BIG-IP system can potentially add 200 bytes of unused bandwidth to the burst reservoir. However, because no bursting has occurred yet, the reservoir is already full at the specified 5,000 bytes, thus preventing the BIG-IP system from storing the 200 bytes of unused bandwidth in the reservoir. In this case, the BIG-IP system simply discards the unused bandwidth.

**If traffic climbs to 1,000 bytes per second (equal to the base rate)**
Still no bursting occurs, and there is no unused bandwidth.

**If traffic jumps to 2,500 bytes per second**
For each second that the traffic continues to flow at 2,500 bytes per second, the BIG-IP system empties 1,500 bytes from the burst reservoir (the difference between the traffic flow rate and the base rate). This allows just over three seconds of bursting at this rate before the burst reservoir of 5,000 bytes is depleted. Once the reservoir is depleted, the BIG-IP system reduces the traffic flow rate to the base rate of 1,000 bytes per second, with no bursting allowed.

**If traffic drops back down to 800 bytes per second**
No bursting is necessary, but now the BIG-IP system can add the 200 bytes per second of unused bandwidth back into the burst reservoir because the reservoir is empty. If traffic continues to flow at 800 bytes per second, the burst reservoir becomes fully replenished from 0 to 5,000 bytes in 25 seconds (at a rate of 200 bytes per second). If traffic stops flowing altogether, creating 1,000 bytes per second of unused bandwidth, then the BIG-IP system adds 1,000 bytes per second into the burst reservoir, thus replenishing the reservoir from 0 to 5,000 bytes in only 5 seconds.

### About the direction setting

Using the **Direction** setting, you can apply a rate class to client or server traffic. Thus, you can apply a rate class to traffic going to a client, to a server, or to both client and server. Possible values are **Any**, **Client**, and **Server**. The default value is **Any**.

Specifying direction is useful in cases where the nature of the traffic is directionally-biased. For example, if you offer an FTP service to external clients, you might be more interested in limiting throughput for those clients uploading files to your site than you are for clients downloading files from your site. In this case, you would select Server as the direction for your FTP rate class, because the Server value only applies your throughput restriction to traffic going from the client to the server.

### About the parent class

When you create a rate class, you can use the **Parent Class** setting to specify that the rate class has a parent class. This allows the child rate class to borrow unused bandwidth from the ceiling of the parent class. A child class can borrow unused bandwidth from the ceiling of its parent, but a parent class cannot borrow from a child class.

Borrowing is also not p~~ossible betwen two child classes the same~~ parent ~~class or between two~~ unrelated classes.

A parent class can itself have a parent, provided that you do not create a circular dependency. A *circular dependency* is a relationship where a rate class is a child of itself, directly or indirectly.

If a rate class has a parent class, the child class can take unused bandwidth from the ceiling of the parent class. The process occurs in this way:

- If the rate of traffic flow to which the child class is applied exceeds its base rate, the child class begins to deplete its burst reservoir as described previously.
- If the reservoir is empty (or no burst size is defined for the rate class), then the BIG-IP® system takes unused base-rate bandwidth from the ceiling of the parent class and gives it to the child class.
- If the unused bandwidth from the parent class is depleted, then the child class begins to use the reservoir of the parent class.
- If the reservoir of the parent class is empty (or no burst size is defined for the parent class), then the child class attempts to borrow bandwidth from the parent of the parent class, if the parent class has a parent class.
- This process continues until there is no remaining bandwidth to borrow or there is no parent from which to borrow.

Borrowing only allows the child to extend its burst duration; the child class cannot exceed the ceiling under any circumstance.

*Note: Although the above description uses the term "borrowing", bandwidth that a child class borrows is not paid back to the parent class later, nor is unused bandwidth of a child class returned to its parent class.*

### About shaping policy

This setting specifies a shaping policy that includes customized values for drop policy and queue method. The default value is None.

You can create additional shaping policies using the Traffic Management shell (tmsh).

### About queue method

The **Queue Method** setting determines the method and order in which the BIG-IP® system dequeues packets.

A rate class supports two queue methods:

**Stochastic Fair Queue**
*Stochastic Fair Queueing* (*SFQ*) is a queuing method that queues traffic under a set of many lists, choosing the specific list based on a periodically-changing hash of the connection information. This results in traffic from the same connection always being queued in the same list. SFQ then dequeues traffic from the set of the lists in a round-robin fashion. The overall effect is that fairness of dequeuing is achieved because one high-speed connection cannot monopolize the queue at the expense of slower connections.

**Priority FIFO**
The *Priority FIFO* (*PFIFO*) queuing method queues all traffic under a set of five lists based on the Type of Service (ToS) field of the traffic. Four of the lists correspond to the four possible ToS values (Minimum delay, Maximum throughput, Maximum reliability, and Minimum cost). The fifth list represents traffic with no ToS value. The PFIFO method then processes these five lists in a way that attempts to preserve the meaning of the ToS field as much as possible. For example, a packet with the ToS field set to Minimum cost might yield dequeuing to a packet with the ToS field set to Minimum delay.

### About drop policy

The BIG-IP® system drops packets whenever the specified rate limit is exceeded. A drop policy specifies the way that you want the system to drop packets. The default value is **fred**.

*Note: You cannot use **fred** or **red**, if you select **sfq** for the **Queue Method** setting.*

Possible values are:

**fred**
Specifies that the system uses Flow-based Random Early Detection to determine whether to drop packets, based on the aggressiveness of each flow. If you require flow fairness across the rate class, select **fred**.
**red**
Specifies that the system randomly drops packets.
**tail**
Specifies that the system drops the end of the traffic stream.
You can create additional drop policies using the Traffic Management shell (tmsh).

---

**Have a Question?**  |  **Support and Sales >**

**Follow Us**  🐦  in  f  ▶  dc

---

**About F5**
Corporate Information
Newsroom
Investor Relations
Careers
About AskF5

**Education**
Training
Certification
F5 University
Free Online Training

**F5 Sites**
F5.com
DevCentral
Support Portal
Partner Central
F5 Labs

**Support Tasks**
Read Support Policies
Create Service Request
Leave feedback [+]